

STIC-Biotech/ChemLib

54329

From: Baum, Stuart
Sent: Tuesday, January 29, 2002 6:35 PM
To: STIC-Biotech/ChemLib
Subject: Sequence Search....thank you

Please do a sequence search for:
SEQ ID NO:1
SEQ ID NO:2

Please search these sequences against the protein data base only...thanks

EXAMINER # 78896
DATE:1/29/02

SERIAL NUMBER:08/763704

MAIL BOX ROOM: CM1 9E12
OFFICE ROOM # CM1 9D06

RESULTS FORMAT: PAPER

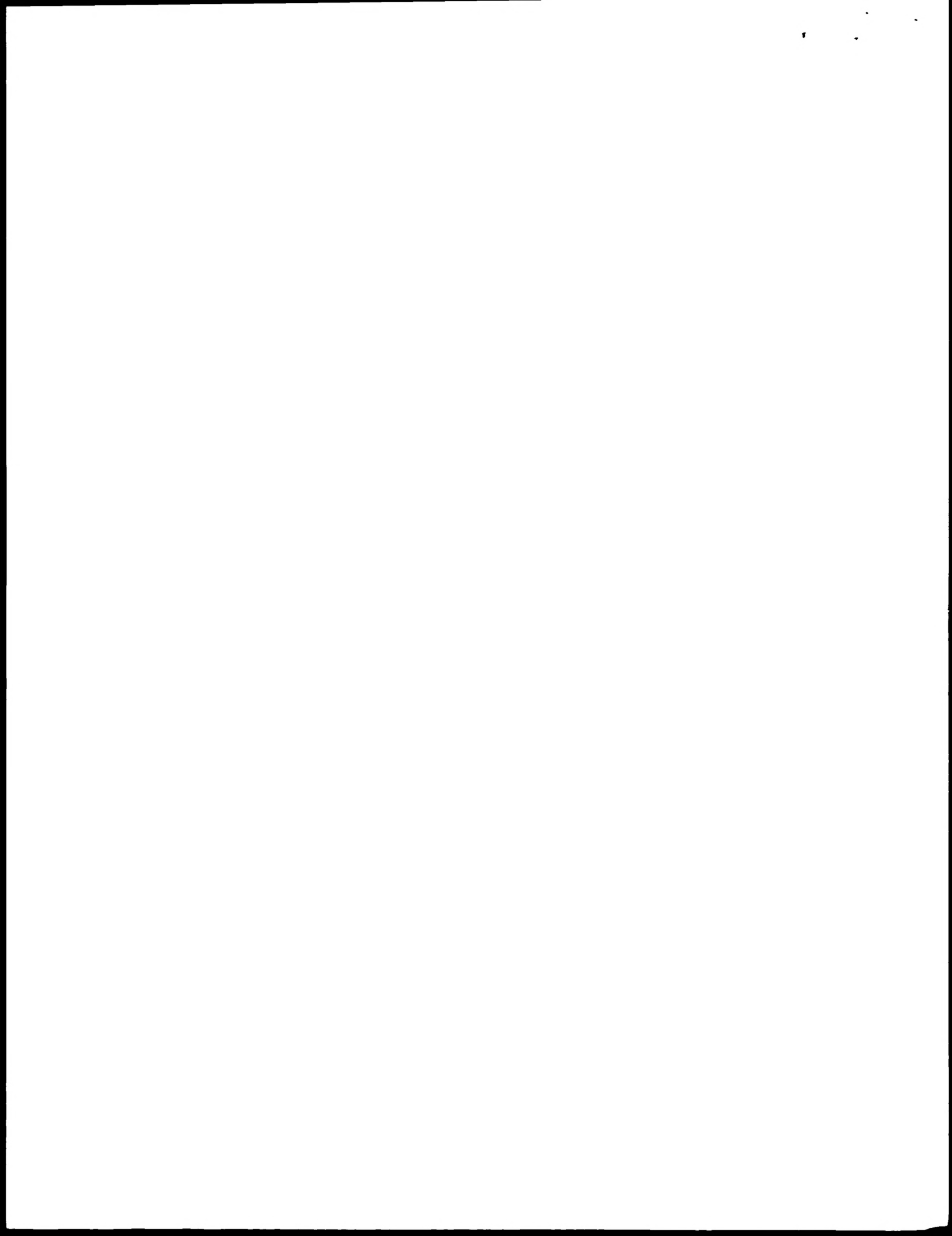
THANKS...Stuart

Stuart F. Baum Ph.D
Art Unit 1638
CM1 9D06
(703) 305-6997
stuart.baum@uspto.gov

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 1/30/02
Date Completed: 1/31/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: 2 - Reverse to
AA Sequences: AA
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 04
WWW/Internet: _____
Other (specify): _____



OM of: US-08-763-704A-1 to: SwissProt_39; * out_format : pfs

Date: Jan 30, 2002 6:58 PM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1994-2000 Compugen Ltd.

Command line parameters:

-MODEL filename: n2p.mon3 BLW xth
-DB SwissProt_39 -COMP fastan -SUFFIX rsp -GAPOP-12.000
-GAPEXT-4.000 -MINMATCH 0.100 -LoopCL 0.000 -LoopEXT-0.000
-GAPAP-4.500 -GAPAPEXT 0.050 -XGAPAP-10.000 -XGAPEXT-0.500
-GAPAP-6.000 -GAPAPEXT 7.000 -GAPAP-10.000 -GAPAPEXT 0.500
-DELOP 6.000 -DELEX 7.000 -START-1 -MATRIX-blosum62
-TRANS human40.cdi -LIST 45 -DEALIGN 200 -THP_SCORE pef
-THP_MAX 100 -THP_MIN 0 -ALIGN 15 -MODE-LOCAL -GAPPEP-fis
-NUM ex1 -HEAPSIZE 63 MINLEN 0 MAXLEN 2000000000
-USER 6326527 -CNCN1 1.63 -NCPU 6 -ICPU 3 -LONGTAXI
-DEV_TIMEOUT 120 -WARN_TIMEOUT 30 -NO_XLDPY -WAIT -THREADS 1

Search information block:

Query: US-08-763-704A-1

Query length: 921

Database: SwissProt_39; *

Database sequences: 160059

Database length: 46664827

Search time (sec): 42.646000

Score list

Accession	Query	Score	EScore	Len	Documentation	
SwissProt_39:ZFA7_MAIZE	1792.00	1792.57	1.74	94	243	P04703 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	1779.00	1797.99	7.40	92	240	P04676 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	1779.00	1794.24	4.36	92	240	P04677 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	1777.00	1733.83	2.80	89	240	P04702 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	815.00	1243.12	6.36	61	234	P04675 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	899.00	1293.73	2.06	61	235	P24449 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	897.00	1229.94	3.26	61	234	P04616 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	893.00	1233.93	7.36	61	234	P04389 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	797.00	1212.31	3.26	60	235	P04704 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	790.00	1203.73	8.40	60	267	P04701 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	790.00	1172.75	5.26	58	240	P04674 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	790.00	1172.20	2.06	57	233	P24470 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	601.00	915.08	1.10	43	240	P04678 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	596.00	771.90	1.40	35	186	P04705 Zea mays (maize)
SwissProt_39:KAF2_SORGH	455.00	753.11	1.10	34	267	P14691 sorghum bicolor (SOR)
SwissProt_39:ZFA7_MAIZE	490.00	743.00	2.06	33	237	P04690 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	470.00	714.19	1.60	32	260	P04699 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	445.00	677.05	1.50	30	269	P04700 Zea mays (maize)
SwissProt_39:ZFA7_MAIZE	445.00	676.14	2.10	30	263	P04679 Zea mays (maize)
SwissProt_39:KAF1_SORGH	440.00	669.04	5.10	30	269	P14690 sorghum bicolor (SOR)
SwissProt_39:ZFA7_MAIZE	457.00	649.24	6.40	29	267	P14692 sorghum bicolor (SOR)
SwissProt_39:ZFA7_MAIZE	265.00	406.97	4.40	15	122	P05815 Zea mays (maize)
SwissProt_39:GLIA_WHEAT	151.00	240.73	2.76	06	355	P13485 tritium aestivum (wl)
SwissProt_39:GH1_WHEAT	131.00	196.10	0.60	10	304	P04729 tritium aestivum (wl)
SwissProt_39:ZFA7_MAIZE	140.00	186.56	0.00	13	790	P24856 notonhea coriopsis (wl)
SwissProt_39:GH1_WHEAT	124.00	185.87	0.00	39	285	P02893 tritium aestivum (wl)
SwissProt_39:GH1_WHEAT	124.00	185.64	0.00	39	295	P16315 tritium aestivum (wl)
SwissProt_39:ZFA7_WHEAT	131.00	181.40	0.00	70	382	P04723 tritium aestivum (wl)
SwissProt_39:GLA5_WHEAT	116.00	172.82	0.01	86	319	P04725 tritium aestivum (wl)
SwissProt_39:ZFA7_MAIZE	114.00	167.11	0.01	92	677	P04334 homo sapiens (human)
SwissProt_39:ZFA7_MAIZE	114.00	160.53	0.01	78	1599	P25663 saccharomyces cerevisiae
SwissProt_39:GH1_WHEAT	114.00	170.06	0.02	75	307	P10386 tritium aestivum (wl)
SwissProt_39:SNF5_YEAST	114.00	161.95	0.02	68	905	P18480 saccharomyces cerevisiae
SwissProt_39:GPA7_WHEAT	114.00	169.15	0.03	83	413	P04727 tritium aestivum (wl)
SwissProt_39:AMVH_YEAST	112.00	155.66	0.03	92	1467	P09640 saccharomyces cerevisiae
SwissProt_39:ZFA7_MAIZE	112.00	160.71	0.04	71	640	P09642 saccharomyces cerevisiae
SwissProt_39:ZFA7_MAIZE	112.00	158.16	0.04	37	955	P08085 mus musculus (mouse)
SwissProt_39:INPH_HUMAN	111.00	160.88	0.04	87	562	P52594 homo sapiens (human)
SwissProt_39:GPA1_WHEAT	111.00	165.92	0.05	48	262	P04721 tritium aestivum (wl)
SwissProt_39:GPA7_WHEAT	110.00	163.19	0.06	64	407	P16573 tritium aestivum (wl)

SwissProt_39:RDX1_YEAST + 110.00 161.81 0.0661
SwissProt_39:RDX1_YEAST + 109.50 163.57 0.0745
SwissProt_39:RDX1_YEAST + 109.00 165.64 0.0819
SwissProt_39:RDX1_YEAST + 108.50 148.66 0.0965
SwissProt_39:RDX1_YEAST + 108.00 160.41 0.0984
seq_name: SwissProt_39:ZFA7_MAIZE

seq_documentation_block:
1. ZFA7_MAIZE: Zea mays (maize)
AC P04703

13-AUG-1987 (rel. 05, created)

13-AUG-1987 (rel. 05, last sequence update)

01-JUN-1984 (rel. 29, last annotation update)

ZEA1-ALPHA-1-PROTEIN (15 KDa) (GENE A265)

Zea mays (maize)

Brassicaceae, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Triticaceae

Polyploid; Andropogoneae; Zea

NCBI_TaxID=4577

11

SEQUENCE FROM N.A.

REVIEWED: 04217881, PubMed 6897917

Gerardby D.F., Messing J.L., Rubenstein L.

"Sequence analysis and comparison of cDNAs of the Zea mays gene

family."

EMBL J. 1:1329-1335(1982)

11- FUNCTION: ZEINS ARE MAJOR SEED STORAGE PROTEINS.

11- MISCELLANEOUS: THE ALPHA ZEINS OF 15 KDa AND 22 KDa ARE THE

70% OF THE TOTAL ZEIN FRACTION. THEY ARE ENCODED BY A LARG

MULTIGENE FAMILY.

11- REFERENCES: 1. STEINBERG, J. AND J. L. RUBINSTEIN. 1982. J. MOL. BIOL. 157: 1-11.

11- REMARKS: 1. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 2. THE

11- REMARKS: 3. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 4. THE

11- REMARKS: 5. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 6. THE

11- REMARKS: 7. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 8. THE

11- REMARKS: 9. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 10. THE

11- REMARKS: 11. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 12. THE

11- REMARKS: 13. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 14. THE

11- REMARKS: 15. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 16. THE

11- REMARKS: 17. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 18. THE

11- REMARKS: 19. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 20. THE

11- REMARKS: 21. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 22. THE

11- REMARKS: 23. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 24. THE

11- REMARKS: 25. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 26. THE

11- REMARKS: 27. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 28. THE

11- REMARKS: 29. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 30. THE

11- REMARKS: 31. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 32. THE

11- REMARKS: 33. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 34. THE

11- REMARKS: 35. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 36. THE

11- REMARKS: 37. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 38. THE

11- REMARKS: 39. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 40. THE

11- REMARKS: 41. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 42. THE

11- REMARKS: 43. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 44. THE

11- REMARKS: 45. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 46. THE

11- REMARKS: 47. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 48. THE

11- REMARKS: 49. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 50. THE

11- REMARKS: 51. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 52. THE

11- REMARKS: 53. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 54. THE

11- REMARKS: 55. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 56. THE

11- REMARKS: 57. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 58. THE

11- REMARKS: 59. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 60. THE

11- REMARKS: 61. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 62. THE

11- REMARKS: 63. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 64. THE

11- REMARKS: 65. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 66. THE

11- REMARKS: 67. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 68. THE

11- REMARKS: 69. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 70. THE

11- REMARKS: 71. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 72. THE

11- REMARKS: 73. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 74. THE

11- REMARKS: 75. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 76. THE

11- REMARKS: 77. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 78. THE

11- REMARKS: 79. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 80. THE

11- REMARKS: 81. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 82. THE

11- REMARKS: 83. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 84. THE

11- REMARKS: 85. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 86. THE

11- REMARKS: 87. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 88. THE

11- REMARKS: 89. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 90. THE

11- REMARKS: 91. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 92. THE

11- REMARKS: 93. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 94. THE

11- REMARKS: 95. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 96. THE

11- REMARKS: 97. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 98. THE

11- REMARKS: 99. THIS ENTRY IS A MULTISUBUNIT PROTEIN. 100. THE

Thu Jan 31 09:02:59 2002

```

DR      Pfam: PF01559; Zein; 1.
EW      Seed storage protein; Repeat; Multisane family; Signal.
FI      SIGNAL 1 21
FT      CHAIN 22 234 ZFN-ALPHA.
SQ      SEQUENCE 244 AA; 25435 MW; A641FE13961136A2 CRC64;

alignment_scores:
  Quality: 815.00 Length: 244
  Ratio: 3.916 Gaps: 3
  Percent Similarity: 84.016 Percent Identity: 72.541

alignment_block:
US-08-763-704A-1 x ZEAC_MAIZE

Align seq 1/1 to: ZEAC_MAIZE from: 1 to: 244

112 AUGGGGACCAAGAUUUUUGCCUUAUGCCUUGGUGUUGUUGUUG 161
1 MetAlaLysLeuPheCysLeuMetLeuLeuGlyLeuSerAlaSe 17
162 UGUUGGUUAGCGGACAAUUGUUGCCUUAUUGCCUUAUUGGU 211
17 rAlaAlaThrAlaThrilePheProGlnCysSerGlnAlaAla 44
212 CCUGUCUUGCCGCGGUAUUGUUGCCUUAUUGCCUUAUUGGU 261
207 nGlnLeuLeuProPheAsnGlnLeuAlaLeuThrAsnProAlaValPhe 224

alignment_scores:
  Quality: 809.50 Length: 244
  Ratio: 3.911 Gaps: 4
  Percent Similarity: 84.490 Percent Identity: 72.654

alignment_block:
US-08-763-704A-1 x ZEAC_MAIZE

Align seq 1/1 to: ZEAC_MAIZE from: 1 to: 244

112 AUGGGGACCAAGAUUUUUGCCUUAUGCCUUGGUGUUGUUG 161
1 MetAlaLysLeuPheCysLeuMetLeuLeuGlyLeuSerAlaSe 17
162 UGUUGGUUAGCGGACAAUUGUUGCCUUAUUGCCUUAUUGGU 211
17 rAlaAlaThrAlaThrilePheProGlnCysSerGlnAlaAla 44
212 CCUGUCUUGCCGCGGUAUUGUUGCCUUAUUGCCUUAUUGGU 261
207 nGlnLeuLeuProPheAsnGlnLeuAlaLeuThrAsnProAlaValPhe 224

```


seq name: swissprot_397ZFA2_MALZF

55-5	ACUUAUACAGUUGAGCGACGACUAGUACUGGCUUACUUCAGCAACAAACAAAC	674
107		
	nleuLeuArgSerGlnLeuAlaAlaValTyr	ProAlaGlnIleP
167		
575	UUUUUCCAUUUUACGCAAUUGGCGGACACACACGCGCGUGUUAUUUUUAG	624
182		
	reLeuProPheAsnGlnLeuAlaAlaLeuAsnSerIleSalAlaTyrValGln	198
625	CAGCAAAATACUACUACCAUUAUGACGACUAGCGUAGCAAGCAAAACGUAUUUUG	674
199		
	GlnGlnGlnLeuLeuLeuProPheSerGlnLeuAlaAlaValSerProAlaAla	216
675	CUUCUUGGACACAGCAACAGUGUGGUGUUCUUCUACACAGCAAGUUUUUAGUA	724

[illegible]

825 CCCCCU 841
 |||||
 265 aleupic 267

2EA3_MA1ZE
STANDARD,
P06674.

01-JAN-1988 (rel. 06, Last sequence update)

LEIN ALPHA PRECURSOR (19 KDA) (CLONE 19A2) (FRAGMENTS) :

Eukaryota: Viridiplantae; Streptophyta; Charophyta; Equisetidae; Poales; Podales; Podocarpaceae; Pinales; Pinaceae; Pinus

PAHCOIdea: And of course, NCBI_TaxID=4577;

SEQUENCE FROM N.A.

Marks M.D., Lindell J.S., LaRkins B.A., "Nucleotide sequence analysis of zein mRNAs from maize endosperm."

J. Biol. Chem. 200:104-120, 1953.

70% OF THE TOTAL ZEIN FRACTION. THEY ARE ENCLOSED IN A LARVAL

! MISCELLANEOUS: STRUCTURALLY, β -K AND β -GLUC AND TOPOLOGICALLY ANTIPARALLEL HELICES CLUSTERED WITHIN

A DISTORTED CYLINDER.

between the Swiss Institute of Bioinformatics and the EMBL consortium.

use by non-profit institutions as long as this conference usage by and for commercial purposes is not removed.

entities requires a license agreement (see note 17).

EMBL: M12142; AAA33525.1; -.

MaizelR; 530'96; " .

pfam; PF01559; Zein; 1. Multigene family; standard

NON_TER
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852

100

FT CHAIN 19 240 ZEIN ALPHA.
SQ SEQUENCE 240 AA: 25042 MW: 54314846CE519D CR664:

alignment_scores:

Quality: 769.50 Length: 241
Ratio: 3.848 Gaps: 4
Percent Similarity: 82.988 Percent Identity: 70.539

alignment_blocks:

US-08-763-704A-1 x ZEAL_MAIZE

Align Seq 1/1 to ZEAL_MAIZE from: 1 to: 240

```

121 AAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 170
||||||| ::::::::::::::::::::::::::::|
1 LyslethecysphtleuMetleuenglyleuSerAlaSerAlaAlaIath 17
171 CGAGAA'AAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 220
::::::::::::::::::::::::::::::::::::|
17 rAlaIathrlethecysphtleuSerAlaSerAlaAlaIathrlethecysphtleu 34
221 CGCAGAAUUAUUGCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 270
||||||| ::::::::::::::::::::::::::::::|
34 roProtyrleuSerProAlaValSerValCysgluAsnProIleleu 50
271 CAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 320
||||||| ::::::::::::::::::::::::::::::|
51 GlpProtyrAlaIleGlnGlnAlaIleAlaAlaGlylleuProleu 67
321 ACCGURHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 370
||||||| ::::::::::::::::::::::::::::::|
67 rProtyrleuGlnGlnIleProSerAlaIleleuGlnIleleuProleuV 84
359 UGAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 408
||||||| ::::::::::::::::::::::::::::::|
84 AluIleleuAlaIleGlnGlnAlaIleAlaAlaGlnIleleuGlnIleleuVal 100
409 CAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 458
||||||| ::::::::::::::::::::::::::::::|
101 ::::::::::::::::::::::::::::::::::|
459 GAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 508
||||||| ::::::::::::::::::::::::::::::|
109 MetIleGlnIleProtyrleuSerAlaIleAlaAlaIleleuGlnIleleuVal 126
509 CAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 558
||||||| ::::::::::::::::::::::::::::::|
126 LetyrleuGlnIleGlnIleleuProtyrleuSerAlaIleAlaIleleuVal 141
559 UGAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 608
||||||| ::::::::::::::::::::::::::::::|
142 ::::::::::ProGlnIleleuProtyrleuSerAlaIleAlaIleleuVal 176
609 CAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 658
||||||| ::::::::::::::::::::::::::::::|
156 rAlaIleAlaIleleuGlnIleleuProtyrleuSerAlaIleAlaIleleuVal 173
659 CAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 708
||||||| ::::::::::::::::::::::::::::::|
173 spValSerProAlaAlaIleleuProtyrleuSerAlaIleAlaIleleuVal 189
709 CAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 758
||||||| ::::::::::::::::::::::::::::::|
190 LeuIleAlaIleleuGlnIleleuProtyrleuSerAlaIleAlaIleleuVal 226
759 GAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCUUGUCACUUGUGUUA 808
||||||| ::::::::::::::::::::::::::::::|
206 UProtyrleuSerProAlaIleleuProtyrleuSerAlaIleAlaIleleuVal 224
809 AAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCACUUGUGUUA 841
||||||| ::::::::::::::::::::::::::::::|
224 rAlaIleAlaIleleuGlnIleleuProtyrleuSerAlaIleAlaIleleuVal 240

```

seq_database: SwissProt_39:ZEAL_MAIZE

seq_documentation_block:

AC ZEAL_MAIZE STANDARD: PRI: 200 AA.
DT 01-MAR-1992 (Rev. 21, Created)
DT 01-MAR-1992 (Rev. 21, Last sequence update)
DT 15-DEC-1998 (Rev. 47, Last annotation update)
DE ZEIN ALPHA PRECURSOR (19 KDa) (PMS2).
GN ZPM52.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Strophophylla; Embryophyta; Charophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OC Panicoideae; Andropogoneae; Zea.
OX NRI_TAXID 4577;
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN CV. A619;
RX MEDLINE:90060774; PubMed 2663514;
RA Quyle T.J.A., Brown J.W.S., Peck G.L.
RT "Analysis of distal flanking regions of maize 19 KDa zein genes."
RI Gene 80:249-257(1989).
RN [2]
RF SEQUENCE FROM N.A.
RA Landridge P., Brown J.W.S., Pinton J.A., Peck G.L.
RT "Expression of zein genes in *Ascaris* in a modified host."
RI Eur. J. Cell Biol. 49:257-264(1989).
CC 1- FUNCTION: ZEINS ARE MAJOR SEED STORAGE PROTEINS.
CC 1- MISCELLANEOUS: THE ALPHA ZEINS OF 19 KDa AND 22 KDa ARE PART OF A
CC 70% OF THE TOTAL ZEIN FRACTION. THEY ARE ENCODED BY A LARGE
CC MULTIGENE FAMILY.
CC 1- MISCELLANEOUS: STEREOISOMERALLY, 24K AND 19K ZEINS ARE COMPOSED OF
CC NINE ADJACENT, TOPOLOGICALLY AND PARALLEL HELICES OF WHICH WITHIN
CC A DISTORTED CYLINDER.
CC
CC THIS SWISS PRO entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not republished
CC modified and this statement is not removed. Users of this database should
CC confirm that they have a license agreement (See http://www.ebi.ac.uk/submit/)
CC or send an email to license@ebi.ac.uk.
CC
DR EMBL: X58700; CAA41543.1; 1.
DR EMBL: X59526; CAA42105.1; Alt term
DR PIR: S15656; S15656.
DR MaizeDB: 58066;
DR InterPro: IPR002530; Zein.
DR Pfam: PF01559; Zein 1.
KW Seed storage protein; Repeat; Multigene family; Zein.
FT SIGNAL 1 21
FT CHAIN 22 243 ZEIN ALPHA
SQ SEQUENCE 240 AA: 25042 MW: 496466146508667 69943

alignment_scores:

Quality: 762.50 Length: 243
Ratio: 3.775 Gaps: 4
Percent Similarity: 82.787 Percent Identity: 68.865

alignment_blocks:

US-08-763-704A-1 x ZEAL_MAIZE
Align Seq 1/1 to ZEAL_MAIZE from: 1 to: 243
112 AAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCACUUGUGUUA 243
||||||| ::::::::::::::::::::::::::::|
1 MetAlaIathrlethecysphtleuMetleuenglyleuSerAlaSerAlaIath
162 UGAGAGAAUHUHUCGUCUUAUUGCGUCUUGUGUCACUUGUGUUA 213
||||||| ::::::::::::::::::::::::::::::|
17 rValAlaIathrlethecysphtleuSerAlaSerAlaIathrlethecysphtleuVal 43

[illegible]


```

1069 TAAATGAGTCAATACATCTGGAGAGCGGTGGCCGTCUUCUACACAGCAA 1118
|||||
151 TPTTSerGlnLeuAlaGlyValSerProAlaThrPheLeuThrGlnPro 167
|||||
1119 TAHHGTHGTCATCCTCATCACATCCGCCTAACAGTGUGGACAGCCUCUU 1168
|||||
168 GlnLeuLeuProPheTyrlrhisAlaIleProAsnAlaGlyThrLeuLe 184
|||||
1169 A'AAATGCTAACAATGTCGCTGCAATGCGAGTGAACUUCGUACAAACCAG 1218
|||||
184 uGlnLeuGlnGlnLeuLeuProPheAsnGlnLeuAlaLeuThrAsnProA 201
|||||
1219 CAGGHTTTAA'CAATCAATGATGAGTGGGCGGCGCTGCUU 1259
|||||
201 LAlaLysTrpGlnGlnProIlellelsclycylalaleuPhe 214
|||||
seq name: sp_plant.Q41884
cd-documentation_block:
Q41884 PRELIMINARY: PRI: 240 AA.
01 NOV-1996 (JREMLrev. 01, created)
01 NOV-1996 (JREMLrev. 01, Last sequence update)
01 JUN-2001 (JREMLrev. 17, Last annotation update)
ZFN (FRAGMENT).
ZFN.
ZFNays (Matzo)
Eukaryota Viridiplantae; Streptophyta; Embryophyta; Tracheostyza;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Pinales; Pinaceae; Pinus clade;
Panicoideae; Andropogoneae, Zea.
NCBI TaxID 4577;
|||
SEQUENCE FROM N.A.
Submitter: S. Chaudhuri, S. Messing, J.-F. Meunier, J. Drenth, R. H.
Submitted (MAR-1991) to the EMBL Sequence Bank.
EMBL: M60847; AAA34540.1;
Model: 11458; Zscore: 1410; 11458.
InterPro: IPR002530; Zfn.
Pfam: PF01559; Zfn. 1.
NUMBER 240 AA. 240.
SEQUENCE 240 AA: 2615 MW: EPDHB2224005044 CR64:
alignment_scores:
Quality: 776.50 Length: 269
Ratio: 3.698 Gaps: 5
Percent Similarity: 78.067 percent identity: 64.684
alignment_block:
US-08-763704A-2 x Q41884
Align seq 1/1 to: Q41884 from: 1 to: 240
459 AUGGAGAGCAAAAUAHUUUGUUGAUUAUGUCUUCUUGUHUUUUGUAGAAG 508
|||||
1 MetAlaAlaLysIlePheSerIleLeuMetLeuLeuAlaLeuSerAlaCy 17
|||||
509 UGACUGUUAAGGCGAGCAUUUUGCTGGAUGGTCACAGCCUUCUACUAGCUU 558
|||
17 sValLeuAspAlaThrIlePheProGlnIleTySerGlnAlaProIleAlaA 34
|||||
559 CGUHUHUCCCAACUACUUCUCCACACAGCAUGUCUUCAGUUGUCUAGAAAAU 608
|||||
34 IalaLeuLeuProPheTyrlrhisAlaIleProAsnAlaGlyThrLeuLe 50
|||||
609 C'AAAHUHUUCUACUUA'CAAGAUUCAACAGTCAACAGCAGGAGGCAUCUU 658
|||
51 ProThrLeuSerIleProTyrlrhisAlaLeuGlnGlnAlaIleAlaThrSerAsnile 67
|||||
659 ACCUHUHUACAGCTUUKUUGUCUUC'CAACCAUUCAGUACGCTUUAUUCACAGCAU 708
|||||

```


88 -pProPheAsnGlnLeuSerThrLeuAsnIleuAlaAlaTyrLeu..._GlnG 104

955 AA^AACAA^AUA^UACCAACA^AG^UAGUGUGCGCGUAC.....CCCGGG 998
|||||
104 GngInGlnLeuProPheSerGlnLeuAlaThrAlaTyrSerGlnGlnGln 120
|||||
999 CAAUUGUGUGCGAUGGAA^AA^CAGAG^AG^AC^AGCACUCUG^AUG^GUAAUGU 104H
|||||
121 GlnPheLeuProPheAsnGlnLeuAlaAlaLeuAsnProSerAlaTyrPh 137
|||||
1049 A^AAC^AAC^AAC^AUA^UACCAAUUGACAG^CAG^UAGUGUGUGCGAUGGCCUG 1098
|||||
117 GngInGlnInIleLeuLeuProPheSerGlnLeuAlaAlaLaserArgA 154
|||||
1099 CUGGCGURRUGACACACAG^AAC^AAUUGUGUGCGUUCUACCGCACACUGAG 1148
|||||
154 LAsSerPheLeuThrGlnGlnLeuLeuProPheTyrGlnGlnPheAla 170
|||||
1149 CUAAAGUGUGGCGACCGUUAU^AAC^UUGCAACA^AAUUGUGUGCGAUGCACCA 1198
|||||
171 AlaAsnProAlaThrLeuLeuGlnLeuGlnGlnLeuLeuProPheValGI 187
|||||
1199 ACUGGCGUUGGA^AAACCG^ACAGAGUGUGUGUACCAACACAGCGCAUUGUGUG 1248
|||||
187 IleaAlaLeuThrHisProAlaAlaLaserTyrGlnGlnProIleIleGlyG 204
|||||
1249 GUGCGUGUGUGU 1259
|||||
204 IYAlaLeuPhe 207
|||||

seq_name: sp_plant_Q06771

seq_id: Q06771 BLOCK: PRT; 266 AA.

AC Q06771:
01-NOV-1996 (TRIMBUREL_01, Created)
01-NOV-1996 (TRIMBUREL_01, Last sequence update)
01-JUN-2001 (TRIMBUREL_17, Last annotation update)
DE ALPHA-COXIN PRECURSOR (ALPHA-PRELAMIN).
GN ALPHA-3b.
OC Coix laachryma-jobi {Jobs Tears}.
OC Fikariyeta; Viridiplantae; Streptophyta, Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Coix.
OC NCBI_TaxID:4565;
EN [1]
RN SEQUENCE FROM R.A.
RC STRAIN CV. AHLAY; TISSUE SEEDLING;
RP MEDLINE 94222476; PMID 8467075;
RA Oshibomi I.M.M., Leite A.; Yunes J.A., Tarzon M.L.P.N.,
RC de Souza Filho G.A.; Artiga P.;
RF "Sequence analysis of 22 kDa p-1 alpha-coxin genes and their
comparison with hemoglobin zeta and kappa genes reveals highly
conserved protein structure and regulatory elements";
PLANT MOL. BIOL., 21:765-778(1993).
RL -1- FUNCTION: SEED STORAGE PROTEIN. SERVES AS A SOURCE OF NITROGEN,
GC CARBON, AND SULFUR FOR THE YOUNG DEVELOPING SEEDLING.
CC -1- SUBCELLULAR LOCATION: PROTEIN BODIES INSIDE VACUOLES.
OC -1- SIMILARITY: TO MAIZE ZETA AND SOYBEAN KAPLINS
DR EMBL: X64114; CAA44827.1;
DR Medrel: 11671; Collab:1411;11671.
DR Interpro: IPRO02540; Zeln.
DR Pfam: PF01559; Zeln; 1.
KW Seed storage protein; Multicene family; Signal; Repeat.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 266 ALPHA-COXIN.
FT DOMAIN 58 257 10 x APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 85 1.
FT REPEAT 86 100 2.
FT REPEAT 101 113 3.
FT REPEAT 114 132 4.
FT REPEAT 133 156 5.
FT REPEAT 157 175 6.

Align seq 1/1 to 082052 from: 1 to: 267

```

11146 GCGCCUAAACGUGGCAACC.....CGUUAUAAAUAGA 1173
      ||| |||||
224 AsnProValValAserAsnProLeuValValAphedLeuValLeu 140
      ||| |||||
1178 ACAAUUGUUGCAUUGGACCAACGUGGUGUUGACAAACAGAGAU 1227
      ||||| |||||
240 nGlnLeuLeuProPheAsnGlnTLeSerLeuMetAsnProValAphSer 1257
      ||||| |||||
1228 ACCAACAAACCAUCAUUGGUGGUGGUGGUGU 1259
      ||||| |||||
257 rpGlnGlnTLeValValGlySerAlaLeuPhe 267
      ||||| |||||

```

```
seq_documentation_block:
  PRFL1MIN
  005420
```

01-NOV-1996 (TREMBLER)

Coix lachryma-jobi (X)
Eukaryota; Viridipla

NCBI_TaxID=4505;

RC STRAIN CV. ADAPT.
RA MEDLINE-93197294

RT amino acid sequence
RT dimensional structure

RP
[2]
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.

RA Ottoboni L.M.M., de
RA de Souza Filho G.A.

RT conserved protein S. 21
RT plant Mol. Biol. 21

DR InterPro, IP002530
DR InterPro, IP002530
DR InterPro, IP002530

F ^T	SIGNAL	F ^T
	CHAIN	22

alignment scores:

Percent Similarity: 60

US-08-763-704A

459 AUGGCCAGCCCAAAAT

509 UGCUGUACGGCG

Downloaded from ascelibrary.org by New York University on 09/09/14. Copyright ASCE, For All Rights Reserved, No part of this document may be reproduced without written permission from ASCE.


```

111 | 1111111111 | 111::111 | 111 | 111 | 111 | 11
143| AsnProAlaIaIaTyLeuInSerGlnLeuAthe...ProCysAsnProLeu 158
920| AAGCGCAUUGAACUCGUACUGCUUAUUUKACGAACAACAAACUACUACCAU 969
    | 1111111111111111111111111111111111111111111
158| uValAla...AsnAlaIaIaTyLeuGlnGlnGlnGlnLeu..... 171
    | ::::::::::::::::::::GlutInfLeuProAlaLeuSer 179
172 | :::::::::::::::::::: 172
1017| CAAGUGGAGCAAUHGAACUCUCAUGCACTUBAGACAAACAACAACUACU 1066
     | 1111111111111111111111111111111111111111111
180| GluLeuAlaValAlaAsnProAsnSerTYrLeuGlnGlnGlnGlnLeule 196
     | 1111111111111111111111111111111111111111111
1067| AACAHACAGCACGACGACGACGCCGCGGCGUGCGUGCUUGUACACAGC 1116
     | 1111111111111111111111111111111111111111111
196| uproPheAsnGlnValAlaValAlaAsnAsnAlaValTYr...GluGlnG 212
     | 1111111111111111111111111111111111111111111
1117| AACAAUHGGCGCGUGCUACAGCCAGAACACUCGCGCUAACGGGCGACGCUC 1166
     | 1111111111111111111111111111111111111111111
212| InHisGlnLeu.....LeuGlnValAsnProLeuAlaIaIaLaphe 225
     | 1111111111111111111111111111111111111111111
1167| UHACAAATGACAA.....CAAUUGCGCGCAUUGGACCACAAUUGGUUUCAC 1210
     | 1111111111111111111111111111111111111111111
226| LeuGlnGlnGlnGlnAraGlnLeuLeuProPheAsnGlnMetSorLeuMe 242
     | 1111111111111111111111111111111111111111111
1211| AAAGCCACACUGUUDUCAGCAACACACCAACCAUAGUGGUGGCGGCGUUD 1259
     | 1111111111111111111111111111111111111111111
242| TAsnProAlaLeuSerTyrGlnGlnGlnGlnGlnGlnGlnValGlyGlyPhe 258
     | 1111111111111111111111111111111111111111111
seq_name: sp_plant_Q07318
seq_documentation_block:
Q07318 PRELIMINARY; prt: 264 AA.
01-NOV-1996 (TREMBLE). 01, created)
01-NOV-1996 (TREMBLE). 01, last sequence update)
01-JUN-2001 (TREMBLE). 17, last annotation update)
ZEIN ALPHA PEPTIDASE (ZSF42).
ZSF42
Zea mays (Maize).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
Panicoideae; Andropogoneae; Zea.
NCBI TaxID 4577;
SEQUENCE FROM N.A.
STRAIN W22; ISSUES LEAF;
MEDLINE 9247493; PubMed-1501150;
Lin C.N., Rubenstein I.;
"Molecular characterization of two types of 22 kilodalton alpha-zelin
genes in a gene cluster in maize.";
MOL. GEN. GENET. 243:244-253(1992).
RT -F- FUNCTION: ZEINS ARE MAJOR SEED STORAGE PROTEINS.
CC - SUBCELLULAR LOCATION: ENDOSPERM PROTEIN BODIES.
CC - DEVELOPMENTAL STAGE: ACCUMULATES DURING KERNEL DEVELOPMENT.
CC - MISCELLANEOUS: THE ALPHA ZELINS OF 19 KTA AND 22 KTA ACCOUNT FOR 70%
OF THE TOTAL ZEIN FRACTION. THEY ARE ENCODED BY A LARGE MULTIGENE
FAMILY.
EMBL: X55724; GAA39255.1; -.
Mandel; 11367; GenBank:J41111367.
Interpro: IPRO02530; zein.
Plant: PF01559; Zein; 1;
Seed storage protein; Multigene family; Signal; Repeat..
FT SIGNAL 1 21
HY SIMILARITY.
FT CHAIN 1 21
ZEIN ALPHA.
FT DOMAIN 58 254
10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 71
1.
FT REPEAT 72 88
2.
FT REPEAT 89 110
3.
FT REPEAT 111 128
4.
FT REPEAT 129 153
5.
```



```
550 ACUGGCUAUCUCACAGCAACAAACAAUUGGCAUUUAAACAAUUGGCGGCGC 599
      :::::
142 AlaAlaTyr.....ProArgLysPheLeuProPheAsnGlnLeuAlaAl 156
      :::::
600 AUGGAAACCCGCGUGGUAUUUUGCAGCAGCAAAAUACUACUACCAUUUAGCC 649
      :::::
156 GluAsnSerAlaAlaTyrLeuGlnGlnGlnGlnLeuLeuProPheSerG 173
      :::::
650 AGCUAGCUAGCAGCAAAACCGUGUGUGUGUGGACACAGCAACAGUUGGCGUG 699
      :::::
174 IleuAlaAspValSerProAlaAlaPheLeuThrGlnGlnGlnLeuLeu 189
      :::::
700 CUGUUGUAACAGCAGGUGUGGCGUUAACCGCGCAACCGUGUUAACACUACA 749
      :::::
190 ProPheTyrLeuIleAlaMetProGlnAlaTyrThrLeuLeuGlnLeuG 206
      :::::
750 ACAAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 799
      :::::
206 GlnIleLeuLeuProPheAsnGlnLeuAlaLeuThrAsnSerThrValPheT 223
      :::::
800 AACAAACAAACAUCAUUGGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 841
      :::::
224 TyrGlnGlnProIleIleTyrGlyAlaLeuPhe 233
```


•


```
alignment_scores:
  Quality: 1021.50
  Ratio: 4.500
  Length: 267
  Gaps: 2
```

US-08 763-704A-2 x 503417 ..

Align seq 1/1 to S03417 from: 1 to 234

459 AUGGACGACAAAUAUUUUGCCUCAAUAUUGCUUUGCAAG 508

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099																																																																																																																																																																																																																																						
1990	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340

1 MetAlaAlaLys11epheCysLeuLeuMet1enLeuLylenSerAlaLe1

558

509 UGUGGUAUGGAGGAUUUCCGCAAGGCUCAAGGCUUAAGCU 558

17 rAAlAtbBrAlpHeBrGlnCvssSrGlnAlaProlleAlas 34

[7] R. A. Aitken, *Algebraic Combinatorics*, Cambridge University Press, 1990.

559 CCCCCCAUACGUCUCACGACGACGUCUUCAGUAUGUGAAAAU 608

[illegible]

34 orIauleuProTyrLeuSerAlaValSerSerValCysGluAsn 50

609 CCAAUUCUUCUACUUAAGGAUCCAAACAGGCAUUGGACGACGGCAUCUU 658

51 ProIleLeuGlnProTyrArgIleGlnGlnAlaIleAlaAlaGlyIleLe 67

[illegible]

659 ACCUUUAUACACCCUUUGUUCCTTCCAAACAUAUCAAGCCCUAUUACAGCAGU 708

67 uProLeuSerProLeuPheLeuGlnGlnSerSerAlaLeuLeu..... 81

709 UACUUUUUGGCAUUUAUUUGGCA'AAAACAUACAGGCGACAAACUACAA 758

..... 19

100

Downloaded from <http://ajphaphapublications.sagepub.com/> at National Archive Publishing Co on May 12, 2015

226 PheAspGlnLeuAlaLeuThrAsnProThrThrLeuLeuGlnGlnProTh 242

Sequence 247 AA:

Sequence 247 AA;

[illegible]


```

998 GCAAUUUCUUCACAUUAAACACUUGCGACCAUUAUACUACUUCAGUUGUAG 1047
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
196 uGlnLeuLeuProPheAsnGlnLeuAlaValSerAsnSerAlaAlaTyrLe 213
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
1048 UACAAACAACAAACAGUACUACCAUUAUACGACGACGACGUGUGAGCCCU 1097
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
214 euGlnLeuArgGlnGlnLeuLeuAsnProLeuAlaValAlaAsnPro 228
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
1098 .....GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 1141
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
229 LeuValAlaThrPheGlnLeuGlnGln 237
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
1142 CACUGGACGUAACGUGUGGACGUGUUAACAACUGGCAAAAUUGUGUGCCAU 1191
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
238 .....GlnGlnLeuLeuProT 243
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
1192 UCGACAAACUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 1241
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
243 yAsnGlnPheSerLeuMetAsnProAlaLeu.....GlnGlnProTLe 257
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
1242 AUUGGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 1259
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
258 ValGlyClyAlaLeuPhe 263
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

seq_name: /US2-GenBank/Genomes/Arabidopsis/AA1908 [W-AAW24977]

seq_documentation_block:

10 AAW24977 standard; protein: 262 AA.

XX AAW24977:

20-JUL-1998 (first entry)

Maize 24 kDa alpha-zelin (t12 gene product).

XX Floury2; t12 gene; alpha-zelin; signal peptide; transgenic plant;

KW seed, essential amino acid; animal feed; maize, rice; wheat;

KW barley; millet; sorghum.

XX Zea mays.

XX Key Location/Qualifiers

FT Peptide 1..21

FT /label= sig_peptide

FT /note= "signal 1"

XX W09802563-AL.

PN 22 JAN 1998.

XX 11-JUL-1997; 97W0-US11723.

XX 17-JUL-1996; 96US-0021833.

XX (PIONEER HI-BRED INT INC.

PA (UYAR-) UNIV ARIZONA STATE.

XX Beach L, Coleman CE, Larkins BA;

XX W01; 1998-110609/10.

XX N-PSDB; AAV09028.

XX Cereal plants containing trans-gene expressing fusion that includes
 PT signal peptide of the t12 maize gene - and protein having high
 PT content of essential amino acids, producing feeds of improved
 PT nutritional value

XX Disclousure; Fig 2; 37pp; English.

XX this polypeptide comprises a 24 kDa alpha-zelin of maize, encoded
 CC by the floury2 (t12) gene (see AAV09028). It includes a 21 amino
 CC acid signal peptide (see AAW24976) that differs from the signal
 CC peptides of other alpha-zelins in having a C-terminal valine residue

CC rather than an alanine residue. This amino acid conversion in the
 CC signal peptide provides an explanation for its retention on the
 CC protein in t12 endosperm and for the phenotypic effects of t12
 CC mutation. The signal peptide targets the alpha-zelin to the lumen
 CC of the rough endoplasmic reticulum. A claimed cereal plant
 CC contains a transgene comprising a first polynucleotide that encodes
 CC the t12 signal peptide and a second polynucleotide that encodes an
 CC agronomically high-value protein. Also new are seeds produced by
 CC the plants, the transgene itself, and a transome that also
 CC includes the t12 promoter. The second polynucleotide preferably
 CC encodes a protein that has a high content of lysine, tryptophan
 CC Thr so that feeds from transformed maize, wheat, rice, barley,
 CC millet or sorghum will have increased contents of these essential
 CC amino acids in their seeds (all claimed).

XX Sequence 262 AA:

alignment_scores:

Quality: 431.50 Length: 404

Ratio: 2.329 Gaps: 13

Percent Similarity: 61.184 Percent identity: 41.118

alignment_block:

us 08 763 704A 2 x AAW23977

Align seq 1/1 to: AAW23977 from: 1 to: 262

```

459 AUGGACGACGACAAAUUUUGGACGUGUAUUUGGUGUGUGUGUGUGUGUGUG 508
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 MetaAlaThrTrpTyrLeuAlaLeuLeuAlaLeuAlaLeuAlaLeuValSer 17
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
509 UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 558
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
17 rAlaThrAsnValPheTyrLeuProGlnPheSerLeuAlaPro.....Sera 44
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
559 GCGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 608
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
33 tAlaTyrProGlnPheLeuProValThrSerMetGlyPheGlnHis 49
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
609 CCAAUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 658
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
50 ProAlaValGlnAlaTyrArgLeuGlnLeuValLeuAlaAlaSerAla... 65
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
659 AUCUUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 708
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
66 .....LeuGlnGlnProTLeAlaGlnLeuGlnGlnGln 74
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
709 UAGGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 758
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
76 InSerLeuAlaHisLeuThrLeuGlnThrTyrLeuAlaThrGlnGlnGln 92
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
753 .....CUACAACACUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 782
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
93 HisPheLeuProSerLeuSerHisLeuAlaValValAsnProValAlaTyr 109
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
788 CUCUCAGCACAACACGAGUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 837
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
109 rLeuGlnGlnGln.....LeuLeuAlaSerAsn..... 118
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
838 GGGCAACAACAACUACAACAACUUGUGUGUGUGUGUGUGUGUGUGUGUGUG 887
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
119 .....ProLeuAlaLeuAlaAsnValAlaThrTyrGln 129
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
888 CAGCAAA.....CAACACUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 928
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
130 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 946
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
926 AUGGACGACGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 950
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
146 tValAsnProAlaValTyrLeuGlnLeuLeuSerSerProLeuAlaVal 163
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:
951 .....CAATAAACAACAACAACAACAACAACAACAACAACAACAACAACA 968
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

163 aGlyAsnAlaProThrTyrLeuGlnGlnInLeuGlnGlnInLeuVal 179
169 .....UUCACGACGACGCU.....GCUGCCUACCCCG 997
180 ProAlaLeuThrHisGlnLeuAlaMetAlaAsnProAlaThrTyrLeuG 196
998 GCAAAUUCUUCGAAUUAACAAACUUCAGAGAAUUAACUUCAGAGAAUUA 1047
196 GlnInLeuLeuProPheAsnGlnLeuAlaValSerAsnSerAlaAlaTyrL 213
1048 UAAUAAACAAACUACUACUACUACUACUACUACUACUACUACUACUACU 1097
213 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 228
1098 GCUACGCUUCUGACACAGAAACUUAUUCUUCGCGUUCUACUUGCAGAC 1147
229 Leu..... 229
1148 GCUUACGCUUCUGACACGCUUCUUCUUCUUCUUCUUCUUCUUCUUC 1197
230 .....ValAlaThrPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 244
1198 AACUUAUUCUACUACUACUACUACUACUACUACUACUACUACUACUAC 1247
244 InPheSerLeuMetAsnProAlaLeu.....GlnGlnProIleValGly 258
1248 GCUACGCUUCUUCU 1259
259 GlyAlaIlePhe 262

```

seq name: 353221-1474, protein: 1474 aa, JP_AA20301AA_AAY82335

seq documentation block:

10 AAY82335 standard; Protein; 250 AA

AC AAY82335;

DT 22-09-2000 (first entry)

DE Metallothionein vector pMTL4 protein sequence.

EW Human; metallothionein; heavy metal removal.

OS Synthetic.

PN JP2000060561-A.

PD 29-FEB-2000.

PF 21-AUG-1998; 98JP-0235879.

PR 21-AUG-1998; 98JP-0235879.

PA (KAWA-) KAWAKI GIUMSU SHINRO HAYASHI

XX WPI; 2500-249676/22.

DR N FSB; AAA08094.

PT New metallothionein polymer used for removal of heavy metals contains

XX metallothioneins connected together by three amino acid residues -

Example 3; Fig 15; 19pp; Japanese

XX The present invention describes a metallothionein polymer in which n
 CC metallothioneins are connected together and the C-terminal amino acid
 CC residue and the N-terminal amino acid residue of the each adjacent
 CC metallothionein are combined by three amino acid residues xaa. The
 CC metallothionein polymer is useful for the removal of heavy metals. The
 CC present sequence represents a protein sequence of a metallothionein
 CC vector, which is used in the exemplification of the present invention.

XX Sequence 250 AA;

alignment_scores:
 Quality: 169.50 Length: 250
 Ratio: 1.646 Gaps: 0
 Percent Similarity: 45.576 Percent Identity: 20.964

alignment_block:

US-08-763-704a-2.rag x AAY82335

Align seq 1/1 to: AAY82335 from: 1 to: 250

```

1183 AATPCTGTCAGTTGTAAGACAGTGTCA.....AGCTTAAAGCT 1146
1184 AATPCTGTCAGTTGTAAGACAGTGTCA.....AGCTTAAAGCT 1146
42 SerCysSerCysSerCysSerCysSerCysSerCysSerCysSerCysSerCys 48
1145 AATGTCAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 1199
48 sTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 64
1098 CAGGCTGTCAAGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 1152
64 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 70
1051 ..... 1051
80 GlySerCysSerCysSerCysSerCysSerCysSerCysSerCysSerCysSerCys 96
1049 TACATTAAGCATGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 1100
96 sCysSerCysCysProValGlyCysAlaGlyCysAlaGlyCysAlaGlyCysAla 114
999 GCGACGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 1050
114 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 120
949 TGTAAATTAAGCATGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 955
124 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 130
914 GTTGAATTAAGCATGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 945
129 sTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 134
864 GCGACGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 919
156 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 160
814 ACCAAGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 869
171 AATGTCAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 187
797 ..... 797
187 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 244
796 ..... 796
204 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 240
763 AATGTCAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 714
221 SerCysSerCysSerCysSerCysSerCysSerCysSerCysSerCysSerCys 240
714 AATGTCAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 686
235 TCGTCAAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 244

```




seq_name: chr10_716,089,211-716,089,213, 5A_cmb.peptus-08-415-751-b

: GENERAL INFORMATION;

APPLICANT: GUT, JIRI

in seq 1/1 to: 45-08-415 251-b from: 1 to: 362

in seq 1/1 to: 05-08-45 251-b from: 1 to: 362

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

44 seq 1/1 to: 02-08-41 251-b From: 1 to: 362

719 GGGUAAACCGCGCAACCCUCUUAACAACUACAAACAUUGUUGCCCCUUGUC 768

[illegible]

```

27  UCHGURGGGACAJAUAUGAAACAAUAUAUG.....ACAGCAACAAAG 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1066 AsnGlycProThrAspGluThrValLeValVallysThrProThrTh 1082
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
71  CAACAUAACAAGUGAAUGAGAGAGCAACACAGACAGCAACAAUAGC.... 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1082 ValLeSerSerSerLea.....SerSerSerSerGlyGln 1096
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117  .....GAGCAAGAUUUUUCCGUCUCCUAUUGUCUU 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1096 LeThrSerSerLeThrSerSerArqProLeThrProPheThrPro 1112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147 UGUCUUCUUCAGUGUGUAGCGGCAAGAAUUAUUCUUAUUGUCUAC 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1114 SerAsnGlyThrSerValLeSerSerSerValLeSerSerSerValTh 1129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 AAGGUCUUAUAGUUCUUGUGUGGCGCAUA..... 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1129 rSerSerLeu....PheThrSerSerProValLeSerSerSerValLeS 1145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
228 .CGUUCGUAUUAUAGUUCUUAUGAUGGAAAGCCAGCUCUUCAGCC 275
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1145 crSerSerThrThrSerThrSerThrLePheSerGluSerSerGlySer 1163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
276 CUUAGGUCUACA....AGCAUUGGAGGAGCAAGCAUAUACCUUUAUCAC 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1162 SerValLeProThrSerSerSerThrSerGlySerSerGluSerGluTh 1178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
323 CGUUGUGUUAUAAAGAGGAGGAGGUAUUGUU.....GGUG 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1178 rSerSerAlaGlySerValSerSerSerSerSerSerSerSerSerSer 1195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 CAGUCUUGGUAACAAUUAUAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1195 crGlySerProThrThrSerSerSerSerSerSerSerSerSerSerAla 1211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
411 ACUGUGUAUACCAAAUAGUGUGGAGAAACUGUUCUUGUUCUUCAGCC 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1212 rThrSerGlnGluThrAlaSerSerSerSerSerSerSerSerSerSer 1228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
461 AACA.....ACAAUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUC 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1228 sThrSerGluGlnThrThrLeuValThrValThrSerGlySerHisV 1245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 UCUAACAUGAA.....ACUGUGUGUGUUAUUGGAGCAACAACU... 527
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1245 aICysThrGluSerLeSerProAlaLeValSerThrAlaThrValThr 1261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
528 .....AUUACAGUACAGCAUUAUAGUUGUUCUUCUUCUUCUUCUUC 568
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1262 ValSerGlyValThrThrGluThrThrThrGlyThrThrThrGlyThr 1278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
569 AACAAUUCUUCUUAUUAACAAUUGUGGAGGAGGAGGAGGAGGAGGAG 618
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1278 rGlnThrThrGlySer...ThrGlySerThrGln...GlnThrThrG 1293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
619 UUGAGCAAGCAAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 668
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1293 LuThrThrGlySerGlnThrValValThrValLeSerSerSerGlySer 1309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
669 UCUUCUUCUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 718
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1310 .....Valcys 1311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
719 GAGCUAAAGCGCAAGGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 768
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1311 sSerGlyThrAlaSerProAlaLeValSerThrSerThrAlaThrLe 1327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
769 CAACUUGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 818
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1328 .....AsnGly...ValThrThrGluThrThrThr 1326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

819 UGUUGCCUUDU 830
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1337 TrpCysProIle 1340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

seq_name: c:\na_data\771aa\EA_NMB pep:US-07-906-349A-6

seq_documentation_block:
: Sequence 6, Application US/07906349A
: Patent No. 5434064
: GENERAL INFORMATION:
: APPLICANT: Schlensing, Joseph
: APPLICANT: Skolnik, Edward Y.
: APPLICANT: Margolis, Benjamin L.
: TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
: TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Browdy and Neimark
: STREET: 419 Seventh Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER AVAILABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07906349A
: FILING DATE: 30 JUN 1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/643,247
: FILING DATE: 18 JAN 1991
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 801 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-906-349A 6

alignment_scores:
   Quality: 108.00      Length: 150
   Ratio: 1.770        Gaps: 9
   Percent Similarity: 40.667      Percent Identity: 23.333

alignment_block:
US-08-763-704A-1/rev x US-07-906-349A-6

Align seg 1/1 to: US-07-906-349A-6 from: 1 to: 801
809 TCTTGTGTACAGAGGCGCTGGCTGTCAAAAGCAAGTTGGACAAAGG 740
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
403 CysCys3 .....CysAlaGlyThr.....Al 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

759 CAACAACCTCTTGTAGTTGTAAGAGGTTGCGAGGTTAGAGGAAA 710
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 yGlyGlyAlaCysAlaCysThrGly...GlyGlyAlaGlyAlaGly... 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
709 GGTAGAAAGCGACCAACTCTGTCTGTGTGTCAGAAAGGAAGATG 660
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
424 .....ThrGlyAlaAlaAlaCysCys..... 430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
659 CGAGGTAGTGTGCTAAATGTAGTAGTATTCTGTGTAGTAAATAAG 610
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```